# 第４講　選択的スプライシング

次のⅠ，Ⅱの各問に答えよ。なお，**図１－１**，**図１－２**，**図１－６**の中，白い四角部分はエキソンをあらわし，山型の実線はスプライシングにより除去される領域をあらわす。

Ⅰ　次の文章を読み，問Ａ～Ｄに答えよ。

真核細胞において，核内でDNAから(ア)転写されたmRNA前駆体の多くはスプライシングを受ける。(イ)スプライシングが起きる位置や組み合わせは一意に決まっているわけではなく，細胞の種類や状態などによって変化する場合がある。これを選択的スプライシングと呼ぶ。選択的スプライシングは，mRNA前駆体に存在する様々な塩基配列に，近傍のスプライシングを促進したり阻害したりする作用を持つタンパク質が結合することによって，複雑かつ緻密に制御されている。例えば，(ウ)哺乳類の$\alpha$-トロポミオシン遺伝子は，**1a**から**9d**まで多くのエキソンを持つが，発現する部位によって様々なパターンの選択的スプライシングを受け（**図１－１**），これによって作られるタンパク質のポリペプチド鎖の長さやアミノ酸配列も変化する（**表１－１**）。

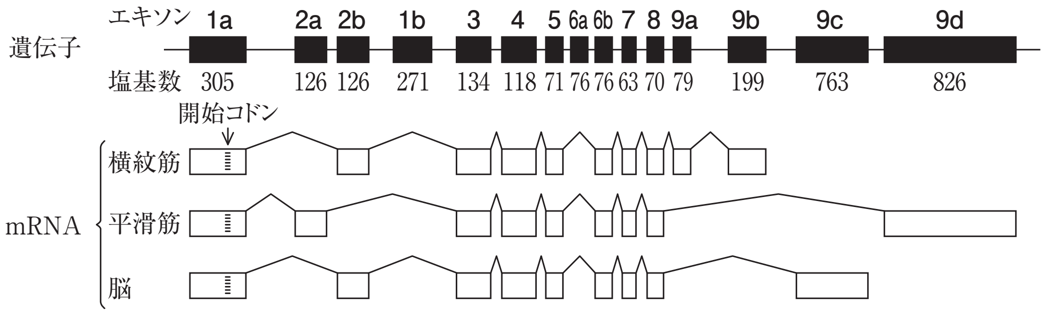


**図１－１**　$\alpha$-トロポミオシン遺伝子の選択的スプライシングの例

**表１－１**　各発現部位における$\alpha$-トロポミオシンタンパク質のポリペプチド鎖の長さ

| 横紋筋 | 平滑筋 | 脳 |
|---|---|---|
| 284アミノ酸 | 284アミノ酸 | 281アミノ酸 |

近年，スプライシングを補正してヒトの遺伝病の治療につなげようとする研究が精力的に行われている。ヒトの５番染色体に存在する *SMN1*(survival motor neuron 1) 遺伝子とそのすぐ隣にある *SMN2* 遺伝子は，塩基配列がほとんど同じであるが，**図１－２**に示す通り，(エ)エキソン７内部のある１つの塩基が，*SMN1* 遺伝子ではＣであるのに対し，*SMN2* 遺伝子ではＴになっているという違いがある。これにより，*SMN2* 遺伝子から作られる mRNA の約９割では，スプライシングの際にエキソン７が使用されず，スキップされた状態となっている。このようにエキソン７がスキップされた mRNA から作られるタンパク質（*Δ*7 型 SMN タンパク質と呼ぶ）は安定性が低く，すぐに分解されてしまう。一方，*SMN2* 遺伝子から作られる mRNA の残りの約１割では，スプライシングの際にエキソン７が使用され，*SMN1* 遺伝子由来のタンパク質と同じアミノ酸配列を持つタンパク質（全長型 SMN タンパク質と呼ぶ）が作られる（**図１－２**）。ヒトにおいて，*SMN1* 遺伝子の欠損を原因とする脊髄性筋萎縮症と呼ばれる遺伝病が知られている。最近，(オ)脊髄性筋萎縮症の治療に，スプライシングを補正する作用を持つ人工的な核酸分子Ｘが有効であることが示され，注目を集めている。

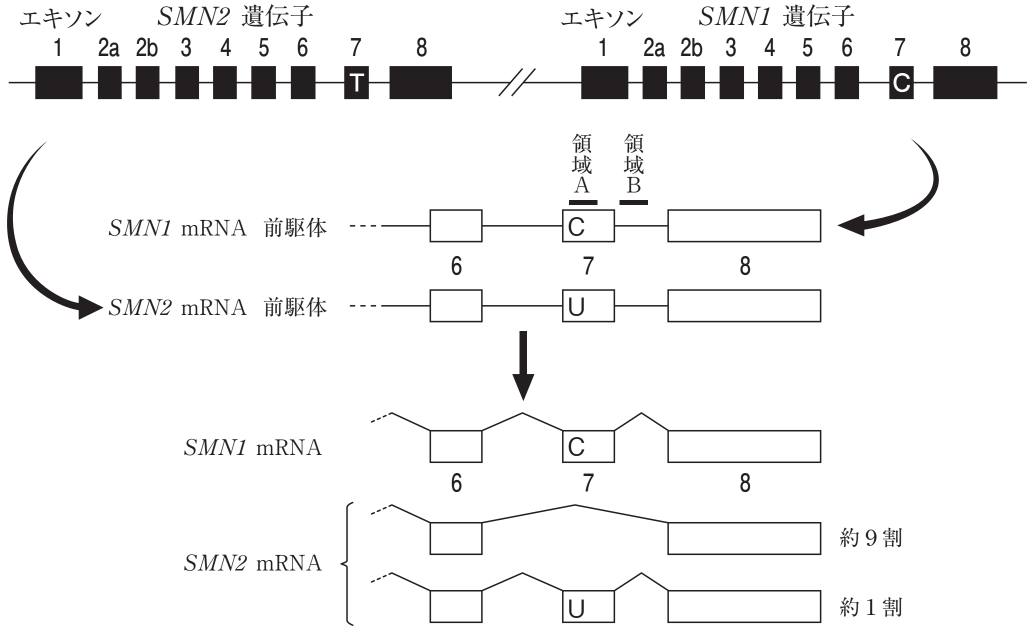


**図１－２**　ヒトの*SMN1*遺伝子と*SMN2*遺伝子およびそれらの転写とスプライシング

〔問〕

A　下線部(ア)について。真核生物における転写の基本的なメカニズムについて，以下の語句をすべて用いて３行（105字）程度で説明せよ。同じ語句を繰り返し使用してもよい。

基本転写因子，プロモーター，RNAポリメラーゼ，片方のDNA鎖，5′→3′

B　下線部(イ)について。異なる塩基配列の６つのエキソン（エキソン１～６と呼ぶ）を持つ遺伝子があるとする。スプライシングの際，エキソン１とエキソン６は必ず使用されるが，エキソン２～５がそれぞれ使用されるかスキップされるかはランダムに決まるとすると，理論上，合計で何種類のmRNAが作られるか答えよ。ただし，スプライシングの際にエキソンの順番は入れ替わらず，エキソンとイントロンの境目の位置は変わらないものとする。

C　下線部(ウ)について。$\alpha$-トロポミオシンmRNAの開始コドンは，**図１－１**に点線で示すとおり，エキソン1aの192～194塩基目に存在する。**図１－１**および**表１－１**の情報から，平滑筋で発現している$\alpha$-トロポミオシンmRNA上の終止コドンは，どのエキソンの何塩基目から何塩基目に存在すると考えられるか答えよ。

解答例：エキソン1bの51～53塩基目

D　下線部(エ)および(オ)について。以下の文中の空欄ａ～ｅに当てはまるもっとも適切な語句を，以下の選択肢①～⑩から選べ。同じ選択肢を繰り返し使用してもよい。

解答例：ａ－①，ｂ－②

*SMN1* mRNA 前駆体の領域Ａ（**図１－２**）の塩基配列はCAGACAAであり，スプライシングの制御に関わるタンパク質Ｙは，この塩基配列を認識して結合する。しかし，*SMN2* mRNA 前駆体の領域Ａの塩基配列は［　ａ　］となっており，ここにはタンパク質Ｙは結合できない。これらのことから，タンパク質Ｙには，スプライシングの際にエキソン７が［　ｂ　］されることを促進するはたらきがあると考えられる。

一方，*SMN1* mRNA 前駆体と *SMN2* mRNA 前駆体で共通の領域Ｂ（**図１－２**）には，スプライシングの制御に関わるタンパク質Ｚが認識して結合する塩基配列が存在する。脊髄性筋萎縮症の治療に有効な人工核酸分子Ｘは，領域Ｂの塩基配列と相補的に結合し，タンパク質Ｚの領域Ｂへの結合を阻害すると考えられている。これらのことから，タンパク質Ｚには，スプライシングの際にエキソン７が［　ｃ　］されることを促進するはたらきがあり，人工核酸分子Ｘは，［　ｄ　］遺伝子のスプライシングを補正することによって，［　ｅ　］型 SMN タンパク質を増加させる作用を持つと考えられる。

① TAGACAA　② CATACAA　③ UAGACAA　④ CAUACAA
⑤ 使用　⑥ スキップ　⑦ *SMN1*　⑧ *SMN2*
⑨ *Δ*７　⑩ 全長

II　次の文章を読み，問Ｅ～Ｈに答えよ。

近年の塩基配列解析装置の急速な進歩によって，生体内に存在するRNAを網羅的に明らかにする「RNA-Seq」と呼ばれる解析を行うことが可能になった（**図１－３**）。例えば，今日用いられているある装置を用いてRNA-Seqを行った場合，長いRNAの塩基配列全体を決定することはできないが，それらのRNAを切断することで得られる短いRNAについて，数千万を超える分子数のRNAの塩基配列を一度に決定することができる。こうして決定される一つ一つの短い塩基配列を「リード配列」と呼び，DNAに含まれる4種類の塩基を表すＡ，Ｃ，Ｇ，Ｔのアルファベットをヌクレオチド鎖の5′→3′の順に並べた文字列として表す（塩基配列決定の際にRNAはDNAに変換されるため，ＵはＴとして読まれる）。リード配列を決定した後，そのリード配列の元となった短いRNAがゲノム中のどの位置から転写されたRNAに由来するかを決めるためには，コンピュータを用いて，ヌクレオチド鎖の向きも含めてリード配列と一致する塩基配列がゲノムの中に出現する位置を見つける「マッピング」と呼ばれる解析を行う。今日の生物学では，このように膨大なデータを情報科学的に解き明かしていくバイオインフォマティクスが重要となっている。

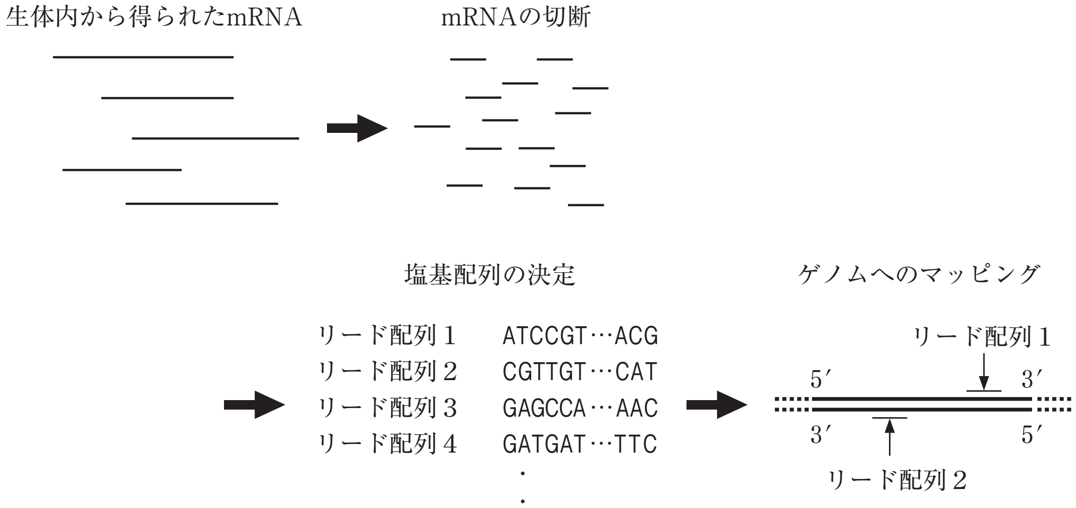


**図１－３**　mRNAを対象としたRNA-Seqの概略図

〔問〕

E　ヒトのゲノム（核相 $n$ の細胞が持つ全 DNA）の塩基対数はおよそ $3\times10^{\boxed{f}}$ である。空欄 f に当てはまる整数を答えよ。

F　一般に，真核生物の遺伝子から転写された mRNA 前駆体には，スプライシングが起きるほか，アデニンが多数連なったポリ A 配列と呼ばれる構造が付加される。これを利用して，真核生物の生体内から得られた RNA を，ある塩基が多数連なった一本鎖の DNA が結合した材質に吸着させることで，mRNA を濃縮して解析することができる。その塩基の名称をカタカナで答えよ。

G　リード配列が「ある特徴」を持つ場合，そのリード配列と一致する塩基配列はゲノムの２つのヌクレオチド鎖の全く同じ位置に出現する（**図１－４**）。「ある特徴」とはどのようなものかを考え，その特徴を持つ 10 塩基の長さの塩基配列の例を１つ答えよ。塩基配列は A，C，G，T のアルファベットを 5′→ 3′の順に並べた文字列として表すものとする。

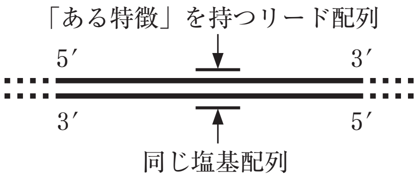


**図１－４**　「ある特徴」を持つリード配列のマッピング

H　真核生物の生体内から得られた mRNA サンプルに対して RNA-Seq を行い，得られたリード配列をゲノムに対してマッピングし，各遺伝子の各エキソン内にマッピングされたリード配列の数を数えた（**図１－５**）。RNA-Seq において mRNA は短い RNA にランダムに切断され，解析装置に取り込まれて塩基配列が決定されたとする。リード配列は各エキソンの長さに比べれば十分に短い一定の長さを持ち，いずれかの遺伝子のエキソン内の１カ所に明確にマッピングされたものとして，以下の問(あ)～(う)に答えよ。

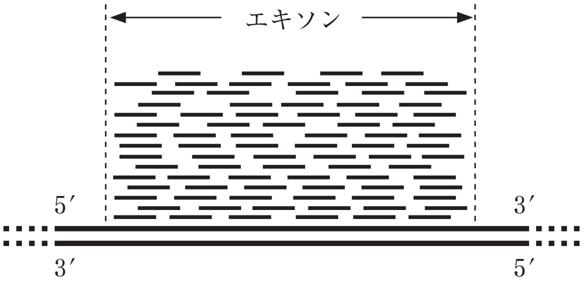


**図１－５**　ある遺伝子のエキソンに多数のリード配列がマッピングされた様子

(あ)　遺伝子１～６のエキソンの塩基数の合計と，エキソン内にマッピングされたリード配列の数の合計は**表１－２**に示すとおりであった。このことから，遺伝子１～６のうち，mRNA の分子数が最も多かったものは遺伝子［ g ］，最も少なかったものは遺伝子［ h ］であったと考えられる。空欄 **g**，**h** に入る数字を答えよ。ただし，遺伝子１～６は選択的スプライシングを受けないものとする。解答例：**g**−1，**h**−2

**表１－２**　RNA-Seq の結果（遺伝子１～６）

| | 遺伝子 1 | 遺伝子 2 | 遺伝子 3 | 遺伝子 4 | 遺伝子 5 | 遺伝子 6 |
|---|---|---|---|---|---|---|
| エキソンの塩基数の合計 | 1000 | 800 | 3000 | 2500 | 1500 | 1800 |
| エキソン内にマッピングされたリード配列の数の合計 | 4500 | 50 | 10000 | 150 | 7000 | 9000 |

(い)　遺伝子７は４つのエキソンを持ち，各エキソンの塩基数と，エキソン内にマッピングされたリード配列の数は**表１－３**に示すとおりであった。遺伝子７は選択的スプライシングを受け，エキソン２かエキソン３のいずれか，あるいは両方がスキップされることがある。**図１－６**に示すように，エキソンが一つもスキップされないmRNAの分子数を$x$，エキソン２のみがスキップされたmRNAの分子数を$y$，エキソン３のみがスキップされたmRNAの分子数を$z$，エキソン２と３の両方がスキップされたmRNAの分子数を$w$とおく。いま$x$が０だったとすると，$y$と$z$と$w$の比はこの順番でどのようになるか，最も簡単な整数比で答えよ。解答例：3：2：5

**表１－３**　RNA-Seqの結果（遺伝子７）

| | エキソン１ | エキソン２ | エキソン３ | エキソン４ |
|---|---|---|---|---|
| エキソンの塩基数 | 800 | 600 | 400 | 1000 |
| エキソン内にマッピングされたリード配列の数 | 16800 | 3600 | 3200 | 21000 |

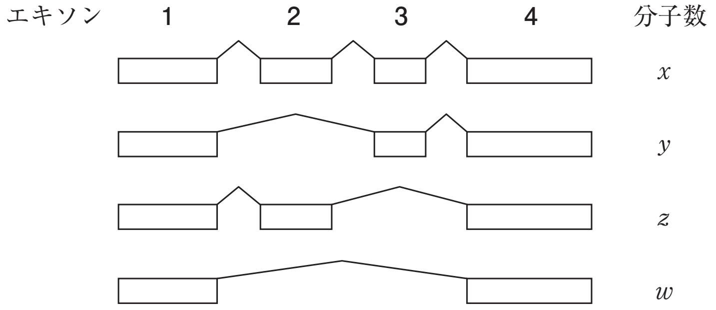


**図１－６**　遺伝子７の選択的スプライシング

(う)　遺伝子７について，$x$が０とは限らないとして，$x$，$y$，$z$，$w$の間に成り立たない可能性がある関係式を以下の選択肢(1)〜(6)から２つ選べ。

(1)　$x < y$　　(2)　$x+z < y+w$　　(3)　$x < w$

(4)　$y > z$　　(5)　$y > w$　　(6)　$z < w$

# 第４講

Ⅰ　A　ある遺伝子の転写が開始される際には，開裂した２本鎖 DNA のうち，片方の DNA 鎖を鋳型として，その遺伝子のプロモーターに基本転写因子と RNA ポリメラーゼが結合し，RNA ポリメラーゼが RNA 鎖を 5′→ 3′の方向へと伸長させる。

B　16 種類

C　エキソン 9d の 81 ～ 83 塩基目

D　a－③，b－⑤，c－⑥，d－⑧，e－⑩

Ⅱ　E　9

F　チミン

G　ATCGATCGAT など

H　(あ)　g－6，h－4

(い)　8：6：7

(う)　(1)，(5)